Larry Thompson
lhthomp@us.ibm.com
January, 2013

Analytics for Logical Dependency Mapping (ALDM)

Increasing pressures to reduce cost and improve resiliency are driving IT organizations to plan and implement complex infrastructure projects

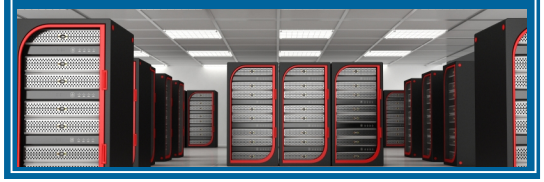
Increasing cost pressures

- 75% of CIOs anticipate a strongly centralized infrastructure in 5 years[1]
- 47% of clients have more than 6 data centers
- 71% of data centers are more than 14 years old
- 58% of clients will be expanding their data centers in next 12-24 months[2]

Resiliency improvement pressures

- 70-90% of manual asset inventories are inaccurate[3]
- 70% of every $1 is spent to maintain and manage the existing infrastructure[4]
- 5-60% of IT workloads may be cloud-enabled[5]
- 14% of CIOs' time is spent removing costs from the technology environment[6]

Infrastructure projects



- **Data center consolidation and relocation**
- **Risk identification and remediation planning**
- **Middleware reclamation**
- **Legacy hardware and software retirement**
- **Server virtualization and consolidation**

Sources
1. IBM Global CIO study, September 2009
2. IBM Global Data Center Study, January 2012
3. IBM estimates, 2012
4. IBM Dynamic Infrastructure client presentation, July 2009
5. IBM research, September 2009
6. IBM Global CIO study, September 2009

ALDM can quickly and unobtrusively discover the IT environment and produce the data and analysis required for IT project planning

Current discovery techniques are insufficient

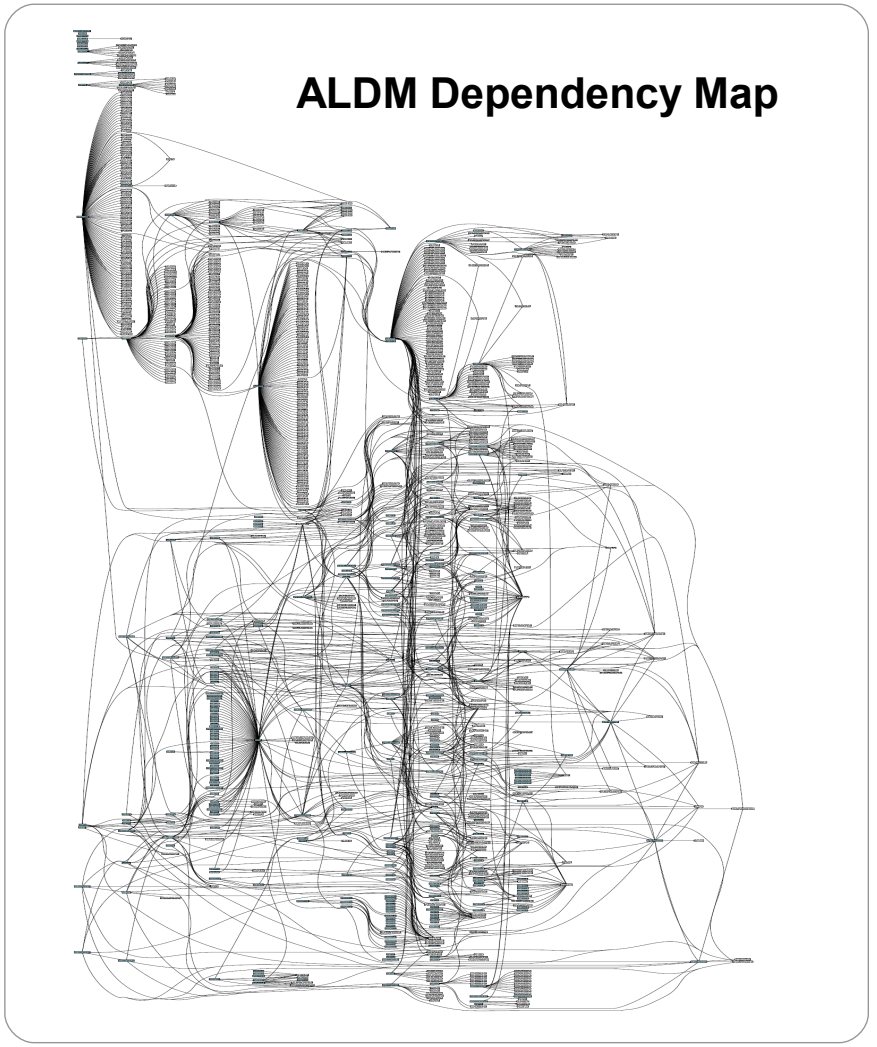
- Current configuration and dependency data are generally less than 70% accurate and complete
- Current tools and manual methods take a long time to produce results, are expensive and generally require credentials.

Quick to deploy and produce results

- A simple script is copied onto each server
- Script requires no credentials, agents or probes
- Script produces data files provided to IBM for post processing
  - Basic server configuration and installed middleware
  - Graphical depictions of server-to-server dependencies
  - Multilevel server dependency analysis
  - Resource utilization trending
  - Dynamic navigation and filtering with drill down to view server-level detail (via Apple iPad).

Developed by IBM Research

- Collaborative effort involved multiple teams in New York, Zurich, Almaden and Moscow.



ALDM can produce results in as little two weeks plus the time needed to implement clients' deployment and change management processes

IBM responsibilities

① Email scripts to client
- Wintel / Linux
- Command line instructions

⑤ Run analytics and produce outputs on the IBM SCE Cloud
- Per server dependency pictures
- Multilevel dependency mapping
- Resource utilization trending
- Dynamic and filtered visualization (iPad)
- .csv with collected data

⑥ Perform analysis
- Application- and user-defined dependencies
- Logical (affinity) groupings for migration
- Infrastructure risks
- IT optimization strategies
- Remediation strategies

Client responsibilities

② Create deployment strategy
- Servers in scope
- Script deployment tool
- .tar file repository
- Change management process

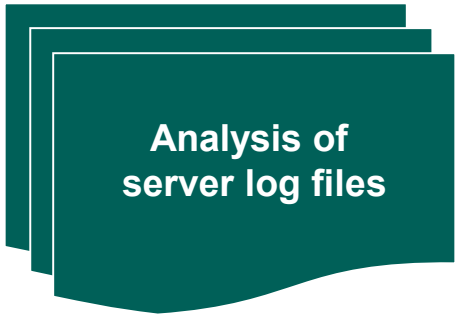
③ Run scripts
- Start date (quiet period)
- Duration (e.g., 5 days)
- Interval (e.g., 15 minutes)

④ Provide input to IBM
- .tar files
- Application-to-server mapping
- Excluded ports
- SFTP to IBM-provided account or physical media transfer

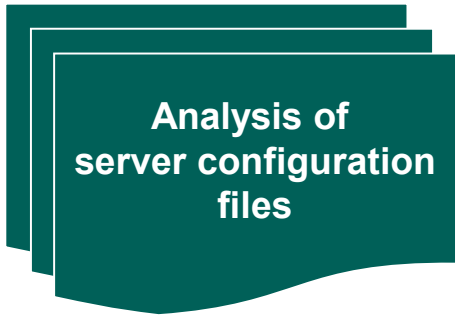
Prepare application-to-server mapping (Dependency chain based on provided applications)

Define ports to exclude, limiting extraneous dependency information

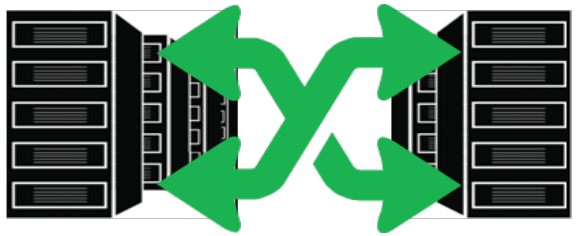
ALDM discovers server configuration information and logical dependencies using three methods*



- Identifies historical dependencies
- Example: connection logs for web servers



- Identifies hardware details and configured dependencies
- Enables ALDM to capture server dependencies that are not observed
- Example: middleware configured to access a database server



Network connections

- Identifies server-to-server activity at specified intervals
- Records observed dependencies by monitoring traffic at each port
- Captures data continuously during the ALDM scanning period, usually 5-7 days

Static data collection

Dynamic data collection

* No application or user data is collected by ALDM

ALDM supports common operating systems and discovers many packaged middleware applications**

Supported operating systems

Windows
- 2000
- 2003
- 2008
- Windows 7

Linux
- Many, see list

VMKernel
- 3.5 and 4.0

AIX
- All

HPUX
- *

Solaris
- 8, 9, 10

Z/OS
- Not supported

Supported middleware*

Database	Version
DB2	6, 7, 8, 9
SQLServer	00, 05, 08
Oracle DB	8, 9, 10, 11
Sybase	15
Teradata	*
Derby/Cloudscape	*
Pointbase	4, 5
PostgreSQL	All
MySQL	All

HTTP Servers	
IBM HTTP Server	All
IBM Proxy Server	All
MS IIS	5, 6, 7
Apache HTTP Server	All

Application servers	
IBM WAS	3, 4, 5, 6, 7
Tomcat	*
WebLogic	8, 9, 10

Messaging	Version
WebSphere MQ	*
WebSphere Message	All
Microsoft Message Q	*

Infrastructure	
MS WNS Server	*
MS DNS Server	*
MS Active Directory	*
Alt-N MDaemon	*
Alt-N RelayFax	*
Symantec AntiVirus	*

Other	
AIX LVM	*
Linux LVM	*
HP-UX LVM	*
Solaris LVM	*

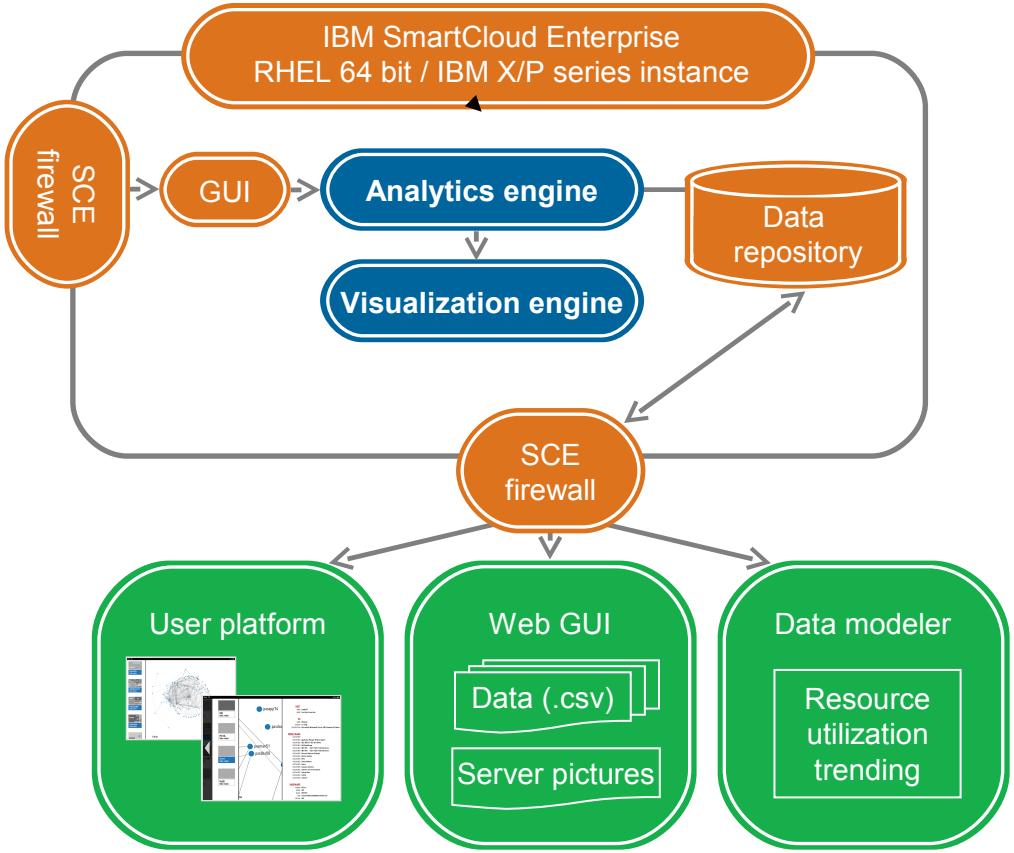
* See your IBM rep for a complete list

** Client-developed applications are not discovered

ALDM runs on IBM SmartCloud Enterprise (SCE) to transform large volumes of complex data into meaningful insights

Infinitely scalable cloud computing platform makes advanced analytics possible

- Database engines for large data volumes
- Easy-to-use GUI for selection of processing options
- Minimum 20x performance improvement over prior systems
  - Processing time shortened from hours to minutes
- iPad-based visualization
  - Dynamically filter and navigate complex data sets for relevant views



The output file contains the data collected and parsed and is useful in understanding the basic environment and identifying potential risks

Top-level server view

HOST_NAME	IP_ADDRESS	OS_NAME	OS_VERSIOI	OS_DISTRIBUTION	PLATFORM	VENDOR	SERIAL
Server_100	255.255.255.1	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	PowerEdge 2900	Dell Inc.	BVG
Server_101	255.255.255.2	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLi		QDKR12G
Server_102	255.255.255.3	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	Pr		NYUS4
Server_103	255.255.255.4	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	P		6TJRR
Server_104	255.255.255.5	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	Pr		N8746FJXSJ
Server_105	255.255.255.6	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProL		CN8746FJXSH
Server_106	255.255.255.7	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant BL460c G1	HP	CN8746FJXSR
Server_107	255.255.255.8	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant BL460c G1	HP	CN8746FJXBB
Server_108	255.255.255.9	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant BL380 G5	HP	CN8746FJXSA
Server_109	255.255.255.10	Windows	5.0.2195	Microsoft Windows 2000 Server	PowerEdge 1600SC	Dell Inc.	NSDO4ASA
Server_110	255.255.255.11	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	PowerEdge SC1420	Dell Inc.	9HDKSJH5
Server_111	255.255.255.12	Linux	2.4.21-27.0.4	Red Hat Enterprise Linux release 3	IBM eServer x226-[87563AA]-	IBM	99KM061
Server_112	255.255.255.13	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant DL380 G5	HP	SKYLK234
Server_113	255.255.255.14	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant BL480c G1	HP	CN784700BP
Server_114	255.255.255.15	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant ML370 G4	HP	SGH857JDS
Server_115	255.255.255.16	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	PowerEdge 2950	Dell Inc.	GFDF44U7SD
Server_116	255.255.255.17	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	PowerEdge 2950	Dell Inc.	SJYDG23SALK
Server_117	255.255.255.18	Windows	5.2.3790	Microsoft(R) Windows(R) Server 2003, Standard Edition	ProLiant DL380 G3	HP	SHGS4765

Unsupported operating system?

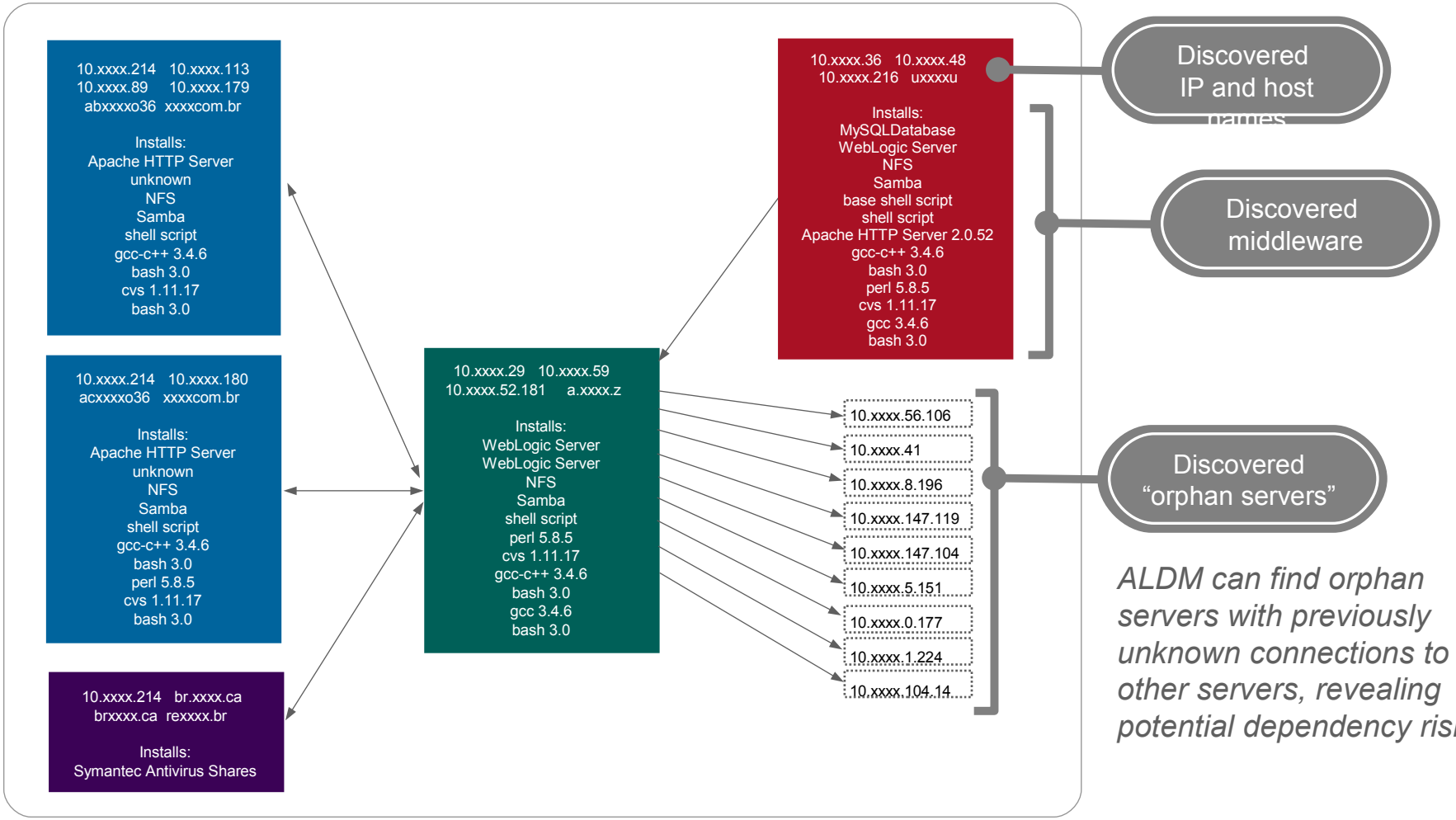
Redundant middleware?

Installed middleware per server

HOST_NAME	MW_CLASS	MW_SUBCLASS	MW_DISTRIBUTION_NAME	MW_VENDOR	MW_VERSION	FYI_MW_INSTALL_PATH
Server_101	DBMS	ORA	Oracle Database	Oracle	9.2	C:\oracle\ora92
Server_101	MAIL	ALNMD	Alt-N MDaemon Email Server	Alt-N		c:\mdaemon
Server_101	FAX	ALNRF	Alt-N RelayFax	Alt-N		c:\relayfax
Server_101	MON	AVSYM	Symantec Antivirus	Symantec		c:\program files\symantec antivirus
Server_101	DBMS	ORA	Oracle Database	Oracle	10.2.0	c:\oracle\product\10.2.0\client_1
Server_101	DBMS	MSQ	Microsoft SQL Server 2005	Microsoft	9.00.4035.00	c:\Program Files\Microsoft SQL Server\
Server_101	DBMS	MYS	MySQL Database	Oracle	5.1	c:\program files\mysql\mysql server 5.1
Server_101	WEB	IIS	Internet Information Services	Microsoft		c:\windows\system32\inetsrv\
Server_101	MON	HPSMH	HP System Management Homepage	HP		c:\hp\hpsmh
Server_101	MON	AVSYM	Symantec Antivirus	Symantec		c:\program files\symantec antivirus

Host names and IP addresses data sanitized

ALDM's graphical server visualization shows directional dependencies and discovered middleware while helping uncover "orphan servers"



ALDM can find orphan servers with previously unknown connections to other servers, revealing potential dependency risks

Multilevel dependency mapping helps automate the process of logically segmenting the infrastructure for data center migration

Multilevel server dependencies

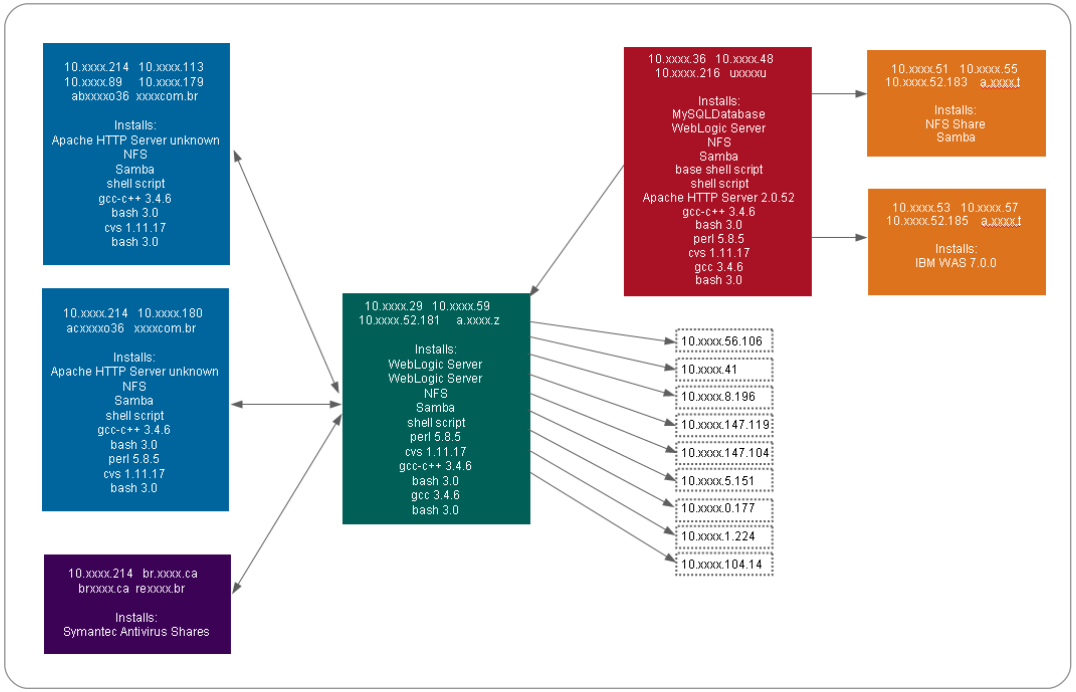
- Maps the end-to-end chain of dependencies
  - Based on user-defined criteria, usually server-to-application mapping
- Automatically creates infrastructure grouping scenarios.

Faster and more accurate than manual grouping techniques

- Simplifies mapping of large numbers of servers, which are too difficult to map manually
- Maps orphan servers, which are sometimes overlooked
- Helps to ensure all dependent infrastructure elements are migrated efficiently to support application needs.

Four levels of mapped dependencies

Web ➔ Application ➔ Database ➔ File Share / WebSphere



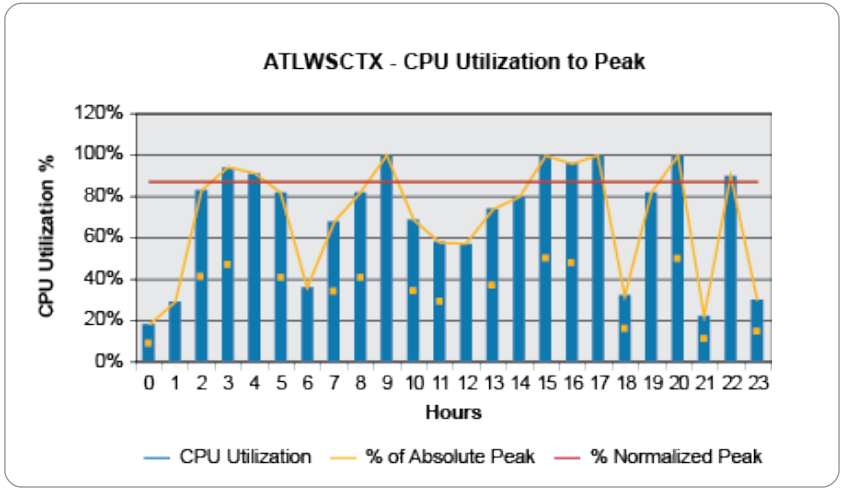
Resource utilization trending helps identify potential servers for optimization, virtualization, consolidation and retirement

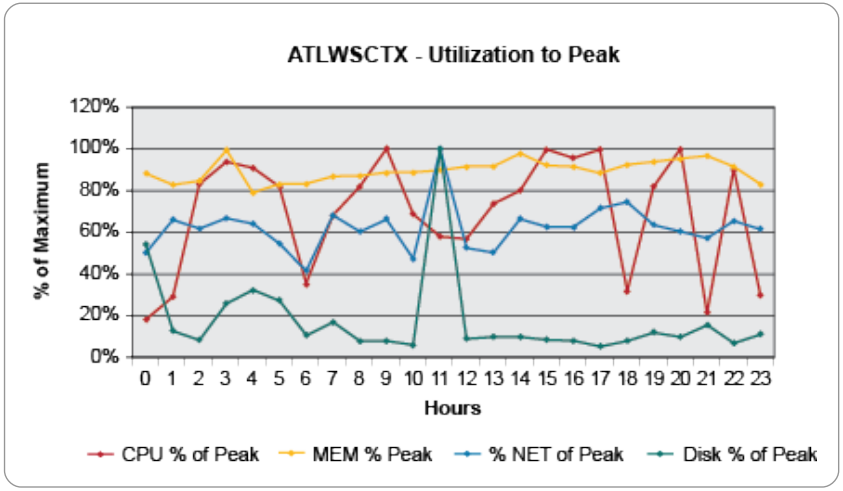
Plots mean and peak utilization for each scanned server

- Shows CPU, memory, network and disk utilization
- Enables users to define desired sampling windows.

Improves resiliency and reduces capital and operating costs

- Identifies best candidates for virtualization, consolidation and retirement
  - Underutilized servers, inconsistently utilized servers, etc.
- Facilitates workload optimization for each server.





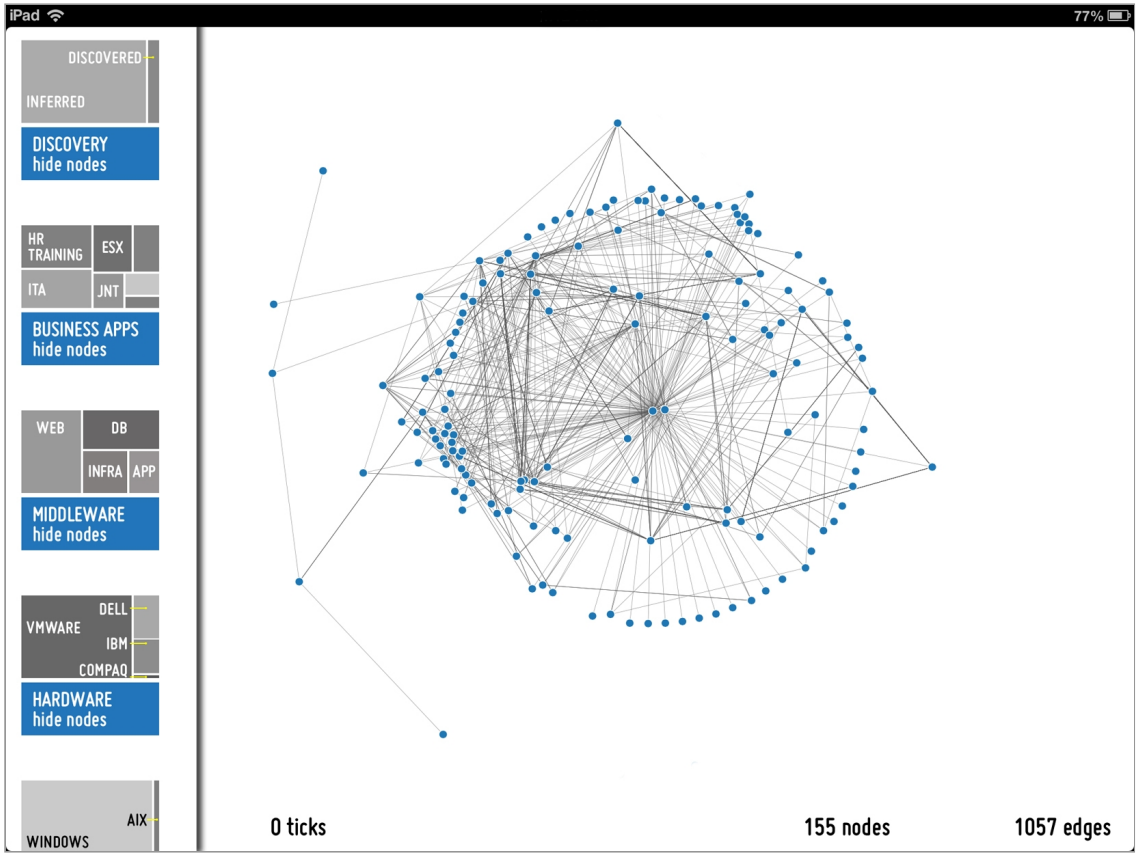
Dynamic visualization and filtering allows for quick and infinitely customized views of even the most complex infrastructures

Dynamic rendering of relevant data

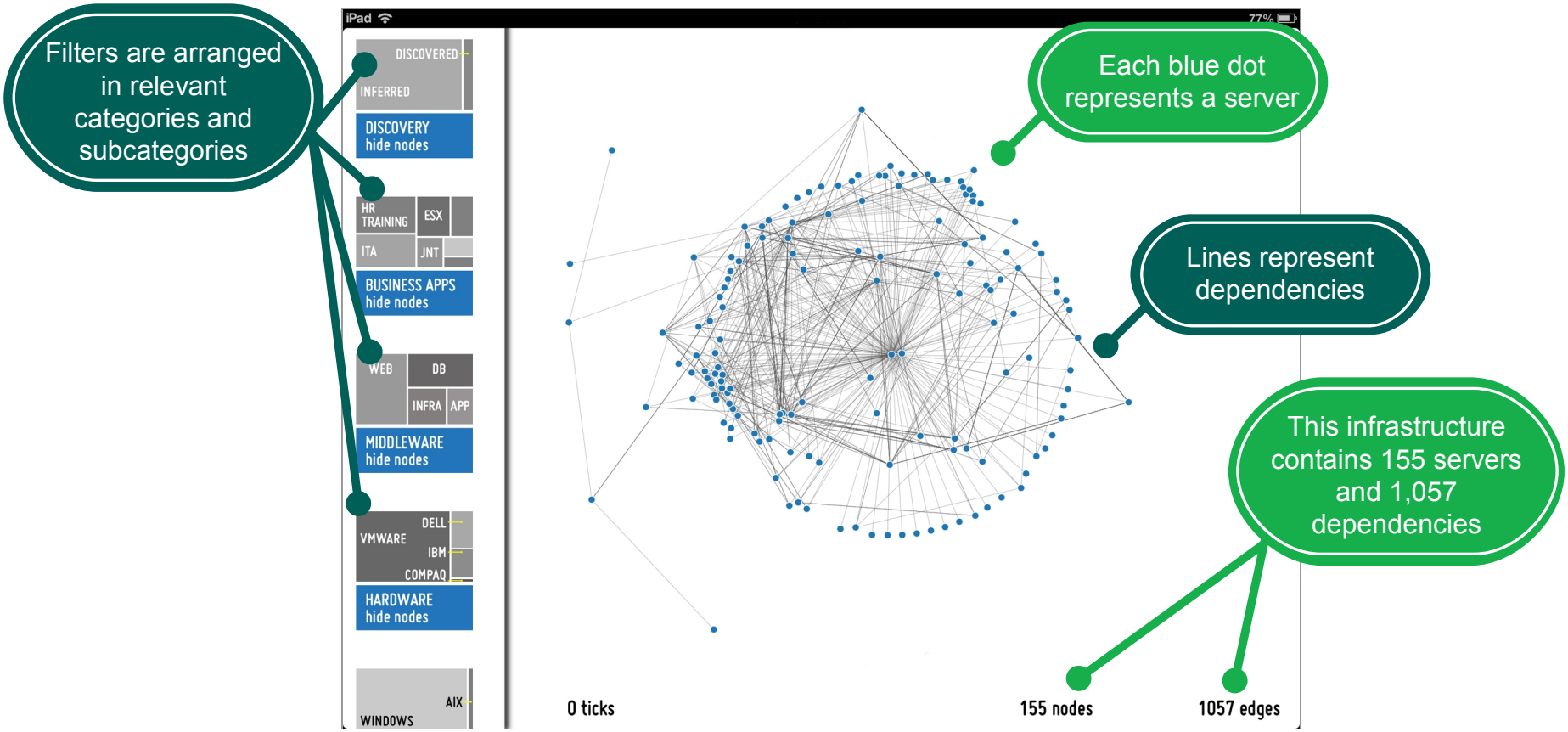
- At-a-glance insights into the infrastructure: number of nodes and degree of dependency
- Global port filtering to reduce extraneous information
- Infinitely customizable filters, enabling users to visualize only data relevant to a specific need
  - Customized filters can be saved for standardized views

Fast and easy navigation

- User-friendly operation on an Apple iPad, using standard iPad navigation techniques
- User-directed drill-down view of server details



IT and migration architects use filters to customize the view, allowing them to view the entire infrastructure or hone in on desired nodes

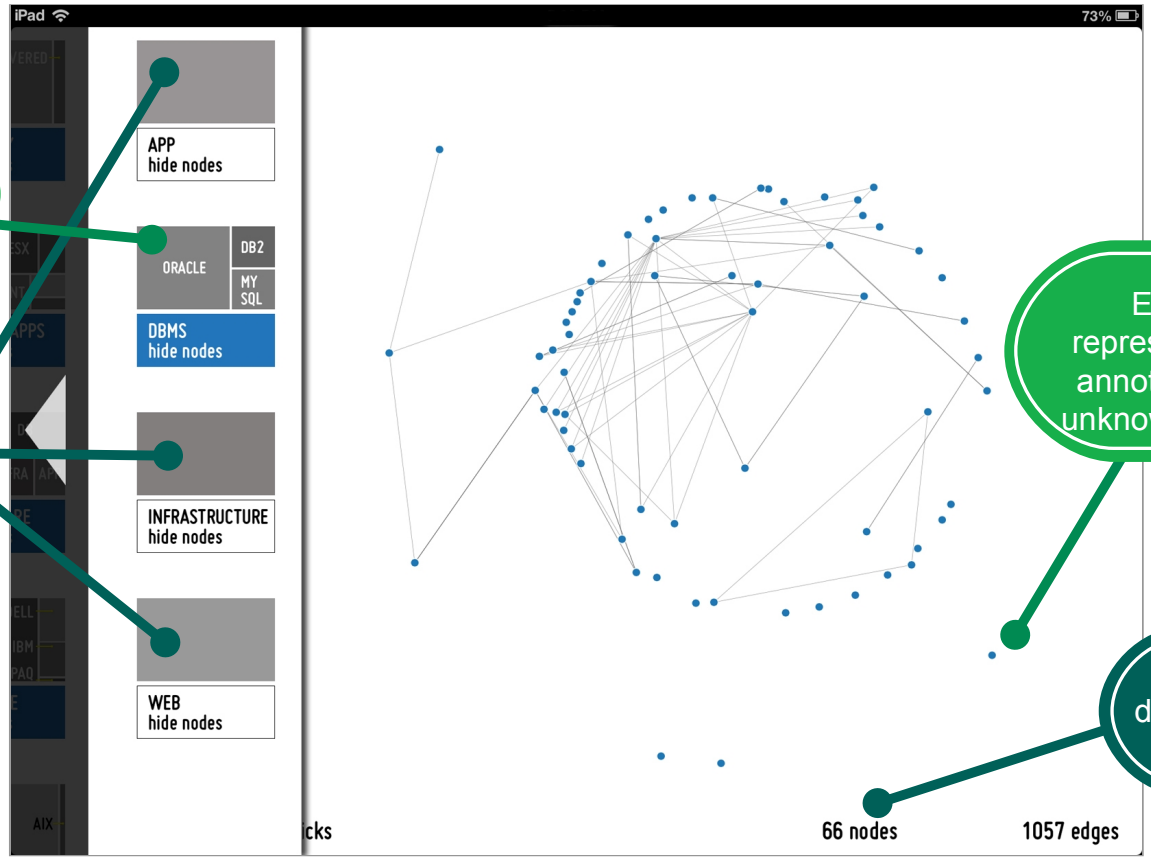


Users tap directly on the filters to view and hide specific server nodes and related dependencies

The dependency map is redrawn as filters are applied and unapplied

Filters can be applied to business applications when clients provide application-to-server maps

In this view only the DBMS servers and dependencies are visualized with all other nodes and dependencies hidden from view



DBMS nodes are visible and active

Web, app and infrastructure nodes are hidden

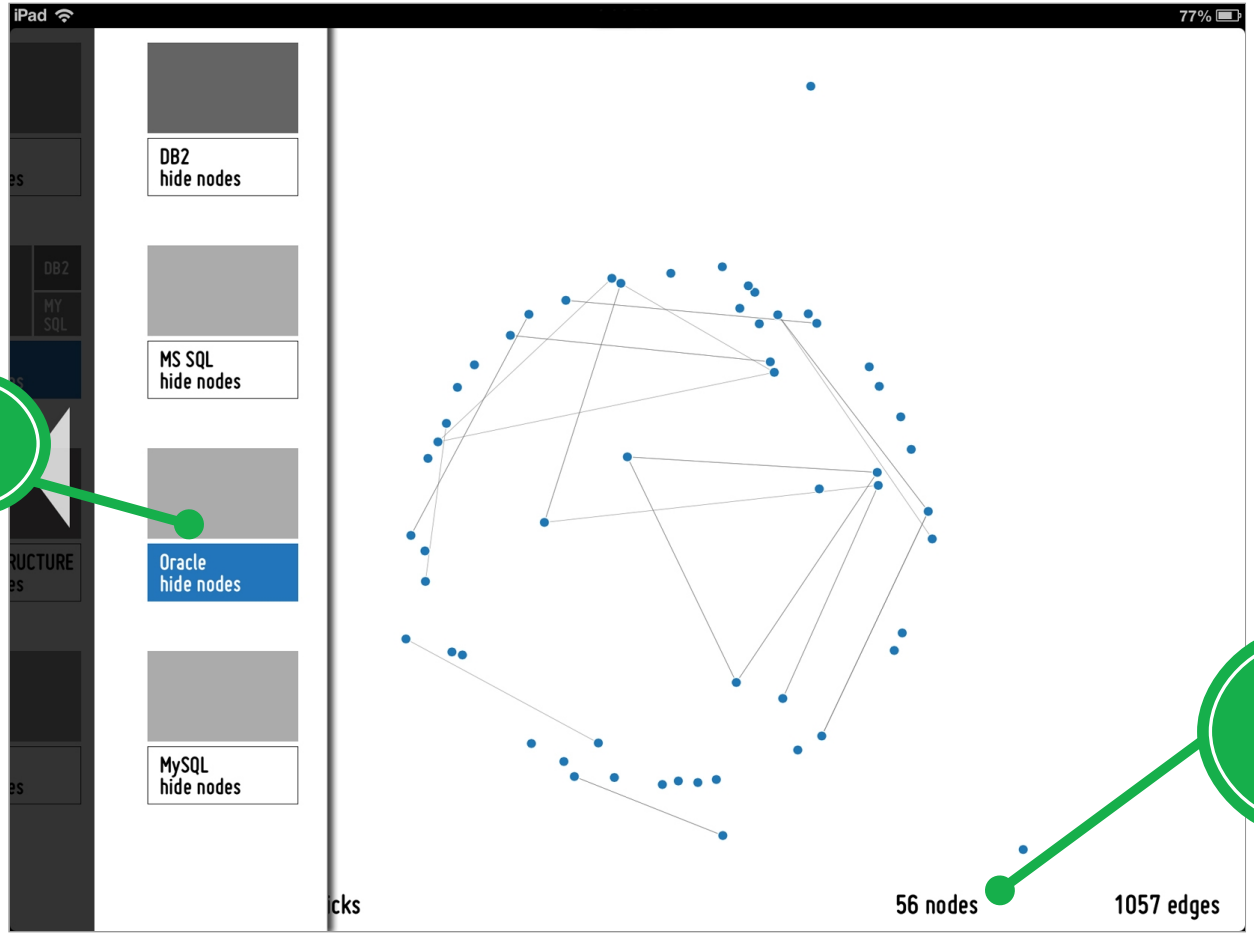
Each lone dot represents a manually annotated server with unknown dependencies

There are 66 database servers in this infrastructure

Multiple levels of filters enable users to visualize specific areas of the infrastructure

Users drill down to the desired level of detail by tapping on the filters

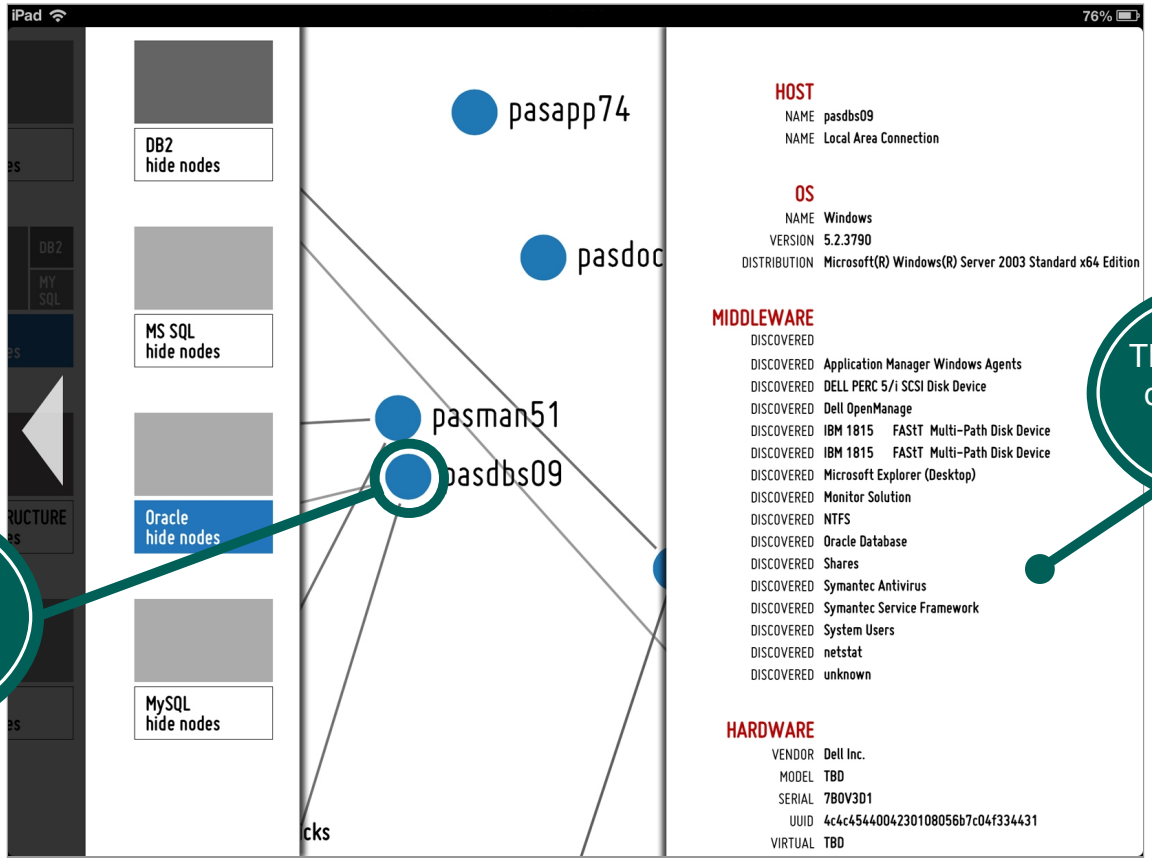
Drilling down further, users can focus on specific database servers, visualizing only Oracle servers and dependencies, for example



Oracle nodes are visible and active

There are 56 Oracle database servers in this infrastructure

Display a servers discovered information by tapping on it



Enlarge the view by pinching fingers outward on the screen, as with other iPad functions

Tap on a node to display detail about it and tap again to hide the detail

Swipe the data panel to scroll up and down and see all available detail for a node

A financial services organization deployed ALDM to better understand its Wintel infrastructure prior to relocating the data center

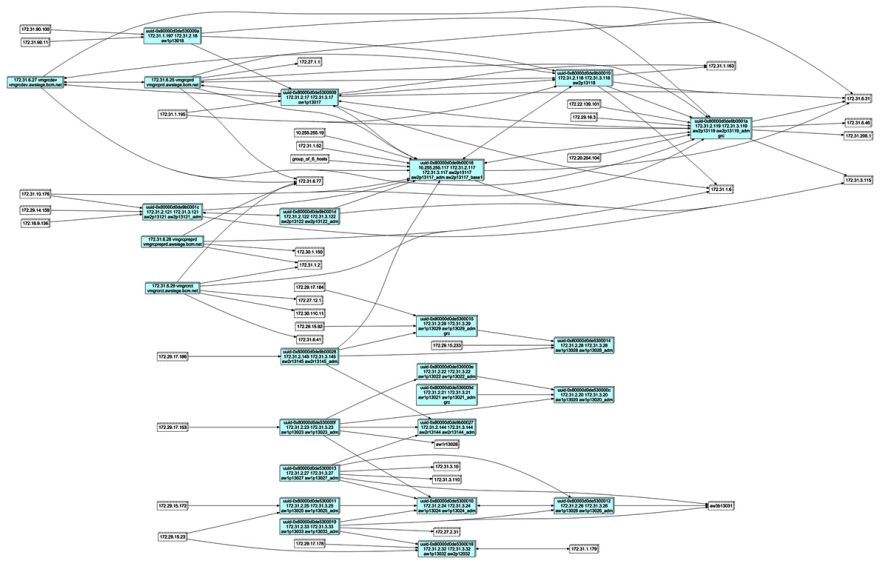
Relocation environment

- 700 Wintel servers
- 2 IBM AIX® midrange system
- 1 IBM System z10™ mainframe
- 400+ critical applications deployed on all platforms

ALDM (1.0) impact

- 50-70% reduction in the time needed to identify dependencies
- Quick generation of affinity grouping scenarios based on how the IT infrastructure actually operates
  - 12 affinity groups defined, one for each major application group
- Seamless transition for end users enabled by a complete view of the infrastructure
- Risk reduction
  - Single point of failure identified and fixed
  - Identification of servers needing OS upgrades

Dependency map for *one* application



Summary

IT discovery has been such a complex task, it is generally not implemented across the entire IT infrastructure. **ALDM lowers the cost and complexity of IT discovery.**

Basic IT discovery is the critical starting point for gathering the data to plan complex infrastructure projects. **ALDM offers a fast and simple process for IT discovering—accelerating it by as much as 30-40 percent.**

Traditional IT discovery methods fail to provide an accurate picture of the IT environment. **ALDM delivers a snapshot of all selected servers and their dependencies**

ALDM output helps IT leaders identify opportunities for cost reduction and resiliency improvement by helping them better understand how the infrastructure actually operates.

Thank you for your time today

For more information

- IBM Site and Facilities Services
- www.ibm.com/services/aldm

Contact

- Larry Thompson, Offering Executive, IBM
- lhthomp@us.ibm.com

Trademarks and notes

IBM Global Services
Route 100
Somers, NY 10589
U.S.A.